RECEIVED

JUL 1 1 2002

1653

TECH CENTER 1600/2900

RAW SEQUENCE LISTING   DATE: 06/21/2002
PATENT APPLICATION: US/09/703,809A   TIME: 11:25:18

Input Set : A:\119941-1083 Seq ID Listing.ST25.txt
Output Set: N:\CRF3\06212002\I703809A.raw

```
 3  <110>  APPLICANT: DeJong, Jeff L.
 5  <120>  TITLE OF INVENTION: Transcription Factors Related to TFIIA
 7  <130>  FILE REFERENCE: 119941-1083
 9  <140>  CURRENT APPLICATION NUMBER: 09/703,809A
10  <141>  CURRENT FILING DATE: 2000-10-30
12  <150>  PRIOR APPLICATION NUMBER: 09/326,529
13  <151>  PRIOR FILING DATE: 1999-06-04
15  <160>  NUMBER OF SEQ ID NOS: 18
17  <170>  SOFTWARE: PatentIn version 3.1
19  <210>  SEQ ID NO: 1
20  <211>  LENGTH: 1617
21  <212>  TYPE: DNA
22  <213>  ORGANISM: Homo sapien
24  <400>  SEQUENCE: 1
25  gctggaggtg ctgtcatggc ctgcctcaac ccggtgccta aactctacag atctgtaatt    60
27  gaagatgtaa ttgaaggagt tcggaatcta tttgctgaag aaggtataga ggaacaagtt   120
29  ttaaaagact tgaagcagct ctgggaaacc aaggttttgc agtctaaagc aacagaagac   180
31  ttcttcagaa atagcatcca atcacctctg tttactcttc agttgccgca cagcttgcac   240
33  caaacattgc aatcgtcaac agcatcatta gttattcctg ctggtagaac tcttccaagt   300
35  tttaccacag cagaactggg cacttcaaac tccagtgcaa actttacttt tcctggttat   360
37  cccattcatg taccagcagg tgtgacacta cagactgtat ctggtcacct ttataaagtc   420
39  aatgtaccaa ttatggtgac agagacttct ggaagagcag gtattcttca gcatccaatt   480
41  cagcaagtat ttcaacagct tggccagcct tcagtaatac aaactagtgt tccacaattg   540
43  aatccatggt ctcttcaagc aactactgaa aaatcacaga gaattgaaac cgtgctacag   600
45  caacccgcaa ttctaccttc tgggccagta gataggaaac acttagaaaa tgccaccagt   660
47  gatatacttg tatctcctgg aaatgagcat aaaaatcgtgc ctgaagcttt gctgtgtcat   720
49  caggaaagtt ctcactatat cagtcttcca ggtgttgtat tttctccaca ggtatctcaa   780
51  acaaattctg atgtggagtc agtgctcagt ggttcagcta gcatggctca aaatctgcat   840
53  gatgagtccc tctccacaag ccctcatggg gctctccact agcacgtgac tgatattcag   900
55  cttcatattc ttaaaaatag gatgtatgga tgtgattctg taaagcaacc aagaaatata   960
57  gaggaaccca gcaacatacc tgtatcagag aaggattcta attctcaggt ggatttaagc  1020
59  attcgggtta ctgatgatga tattggtgaa ataattcaag tagatggaag cggtgataca  1080
61  tcttccaatg aagaaatagg aagtacaaga gatgcagatg agaatgaatt tctaggaat  1140
63  attgacgggg gagatctgaa ggtacctgaa gaagaagctg acagtatttc aaatgaggat  1200
65  tcagccacaa acagtagtga taatgaagac cctcaagtaa acattataga agaggaccct  1260
67  ttaaattctg agatgatgt tagtgaacag gatgtgccag acctgtttga cacggataat  1320
```

     6 21 02

RAW SEQUENCE LISTING   DATE: 06/21/2002
PATENT APPLICATION: US/09/703,809A   TIME: 11:25:18

Input Set : A:\119941-1083 Seq ID Listing.ST25.txt
Output Set: N:\CRF3\06212002\I703809A.raw

```
82 <212> TYPE: PRT
83 <213> ORGANISM: Homo sapien
85 <400> SEQUENCE: 2
87 Met Ala Cys Leu Asn Pro Val Pro Lys Leu Tyr Arg Ser Val Ile Glu
88 1               5                   10                  15
91 Asp Val Ile Glu Gly Val Arg Asn Leu Phe Ala Glu Glu Gly Ile Glu
92              20                  25                  30
95 Glu Gln Val Leu Lys Asp Leu Lys Gln Leu Trp Glu Thr Lys Val Leu
96          35                  40                  45
99 Gln Ser Lys Ala Thr Glu Asp Phe Phe Arg Asn Ser Ile Gln Ser Pro
100       50                  55                  60
103 Leu Phe Thr Leu Gln Leu Pro His Ser Leu His Gln Thr Leu Gln Ser
104 65                  70                  75                  80
107 Ser Thr Ala Ser Leu Val Ile Pro Ala Gly Arg Thr Leu Pro Ser Phe
108              85                  90                  95
111 Thr Thr Ala Glu Leu Gly Thr Ser Asn Ser Ser Ala Asn Phe Thr Phe
112              100                 105                 110
115 Pro Gly Tyr Pro Ile His Val Pro Ala Gly Val Thr Leu Gln Thr Val
116          115                 120                 125
119 Ser Gly His Leu Tyr Lys Val Asn Val Pro Ile Met Val Thr Glu Thr
120       130                 135                 140
123 Ser Gly Arg Ala Gly Ile Leu Gln His Pro Ile Gln Gln Val Phe Gln
124 145                 150                 155                 160
127 Gln Leu Gly Gln Pro Ser Val Ile Gln Thr Ser Val Pro Gln Leu Asn
128              165                 170                 175
131 Pro Trp Ser Leu Gln Ala Thr Thr Glu Lys Ser Gln Arg Ile Glu Thr
132              180                 185                 190
135 Val Leu Gln Gln Pro Ala Ile Leu Pro Ser Gly Pro Val Asp Arg Lys
136          195                 200                 205
139 His Leu Glu Asn Ala Thr Ser Asp Ile Leu Val Ser Pro Gly Asn Glu
140       210                 215                 220
143 His Lys Ile Val Pro Glu Ala Leu Leu Cys His Gln Glu Ser Ser His
144 225                 230                 235                 240
147 Tyr Ile Ser Leu Pro Gly Val Val Phe Ser Pro Gln Val Ser Gln Thr
148              245                 250                 255
151 Asn Ser Asp Val Glu Ser Val Leu Ser Gly Ser Ala Ser Met Ala Gln
152          260                 265                 270
155 Asn Leu His Asp Glu Ser Leu Ser Thr Ser Pro His Gly Ala Leu His
156       275                 280                 285
159 Gln His Val Thr Asp Ile Gln Leu His Ile Leu Lys Asn Arg Met Tyr
160       290                 295                 300
163 Gly Cys Asp Ser Val Lys Gln Pro Arg Asn Ile Glu Glu Pro Ser Asn
164 305                 310                 315                 320
167 Ile Pro Val Ser Glu Lys Asp Ser Asn Ser Gln Val Asp Leu Ser Ile
168              325                 330                 335
                                  Glu Gln Ile Thr Gly Val Arg Gly Ser
```

RAW SEQUENCE LISTING   DATE: 06/21/2002
PATENT APPLICATION: US/09/703,809A   TIME: 11:25:18

Input Set : A:\119941-1083 Seq ID Listing.ST25.txt
Output Set: N:\CRF3\06212002\I703809A.raw

```
179 Glu Asn Glu Phe Leu Gly Asn Ile Asp Gly Gly Asp Leu Lys Val Pro
180     370             375             380
183 Glu Glu Glu Ala Asp Ser Ile Ser Asn Glu Asp Ser Ala Thr Asn Ser
184 385             390             395             400
187 Ser Asp Asn Glu Asp Pro Gln Val Asn Ile Val Glu Glu Asp Pro Leu
188             405             410             415
191 Asn Ser Gly Asp Asp Val Ser Glu Gln Asp Val Pro Asp Leu Phe Asp
192             420             425             430
195 Thr Asp Asn Val Ile Val Cys Gln Tyr Asp Lys Ile His Arg Ser Lys
196         435             440             445
199 Asn Lys Trp Lys Phe Tyr Leu Lys Asp Gly Val Met Cys Phe Gly Gly
200     450             455             460
203 Arg Asp Tyr Val Phe Ala Lys Ala Ile Gly Asp Ala Glu Trp
204 465             470             475
207 <210> SEQ ID NO: 3
208 <211> LENGTH: 3824
209 <212> TYPE: DNA
210 <213> ORGANISM: Homo sapien
212 <400> SEQUENCE: 3
213 ggactttggg actggacaga cctggtcaca gtctaggttc tacatcttac tggtcgagca      60
215 actttagagt caacctattt gatttcttga caagaccaca atctgatccc aaagatgtgc     120
217 tccacaaatc caggcaaatg ggtcaccttt gatgatgatc ctgctgttca atcttctcaa     180
219 aagtcaaaga attttcctct ggagaatcaa ggtgtctgta gaccaaatgg actgaagctg     240
221 aaccctcctg gcctcaggga atttcccagt ggatcttcct ccaccagcag cactcctctc     300
223 tcctccccca ttgtagattt ttatttcagt ccaggacctc caagtaactc tcctctttct     360
225 acacctacca aagacttccc aggttttcct ggcatcccca aagcagggac tcatgtgctt     420
227 tatcctattc cagaatcatc ttcagacagc ccactcgcaa tatcaggagg agaatcttcc     480
229 ttactgccta ccagaccaac atgtttatcc catgccttgt tacccagtga ccactcatgt     540
231 acacatccaa ctcccaaagt aggtcttcca gatgaagtta atcctcaaca ggctgaaagc     600
233 ctaggattcc aaagtgatga tctcccccag tttcagtatt ttcgagagga ctgtgctttt     660
235 tcaagtccat ttcggaaaga tgaaggcagt gattcccatt tcacccttga cccaccagga     720
237 agcaaaaaga tgttctcatc aagaaacaag gagatgccta ttgaccaaaa aagcctaaat     780
239 aagtgttcac tcaactatat ctgtgagaag cttgaacatc tccagtcagc tgagaaccaa     840
241 gactcactta gaagtttgtc tatgcactgt ctatgtgctg aagaaaatgc ctcttccttt     900
243 gtccccaca cactcttcag gagtcagcca aaatccggat ggtctttcat gctgagaatt     960
245 cctgagaaga agaatatgat gtctttccgg caatgggac caattttct gaaagtttta    1020
247 cctggaggaa ttttgcagat gtattatgaa cagggattag aaaatccatt taaatagata    1080
249 cagcttgatc catattgtag gctttctgaa cccaaggttg agaacttcag tttatcagga    1140
251 aaaatccaca ctgtgaagat tgaacatgtg tcttacacag aaaaaaggaa ataccattct    1200
253 aagacagaag tagttcatga acctgacata gagcagatgc tgaagttggg gtccacatcg    1260
255 taccatgact tccttgactt tctgactact gtggaggagg agctgatgaa gttgccagct    1320
257 gtttcaaaac caaaaaagaa ctacgaggag caagaaattt ccttggaaat tgtggacaac    1380
259 ttttggggta aagtcacaaa agaaggaaaa tttgttgaaa gtgctgtgat aactcaaatt    1440
261 tattgcctct gctttgtgaa tgggaacctg gaatgctttt taaccttgaa tgaccttgag    1500
263 ttgccgaagc gagatgaatc ctattatgag aaggactcag aaaaaaaggg gattgatatt    1560
```

RAW SEQUENCE LISTING   DATE: 06/21/2002
PATENT APPLICATION: US/09/703,809A   TIME: 11:25:18

Input Set : A:\119941-1083 Seq ID Listing.ST25.txt
Output Set: N:\CRF3\06212002\I703809A.raw

```
273 aggtcctgtg acaatataag gatacacttt cctgtcccat cgcagtggat caaggccctt  1860
275 tggaccatga acctccagag gcagaagtct ctgaaagcta aaatgaaccg ccgagcatgt  1920
277 ctggggagtt tacaggaact tgaatctgaa cctgtcattc aagtcactgt ggggtcagca  1980
279 aaatatgaga gtgcctacca ggcagtggta tggaagatag atcggcttcc agacaaaaat  2040
281 tcaagtctag atcatcccca ttgtctgtca tacaaattag agcttggatc agaccaagaa  2100
283 attccctctg attggtatcc atttgctact gttcagtttt ccgtgcctga cacctgtgcc  2160
285 tcaaggacag aggtcaggtc tctgggagtg gagagtgatg tccagccaca gaaacatgtt  2220
287 cagcagcgag cttgctacaa catccagcct aaactctaca gatctgtaat tgaagatgta  2280
289 attgaaggag ttcggaatct atttgctgaa gaaggtatag aggaacaagt tttaaaagac  2340
291 ttgaagcagc tctgggaaac caaggttttg cagtctaaag caacagaaga cttcttcaga  2400
293 aatagcatcc aatcacctct gtttactctt cagttgccgc acagcttgca ccaaacattg  2460
295 caatcgtcaa cagcatcatt agttattcct gctggtagaa ctcttccaag ttttaccaca  2520
297 gcagaactgg gcacttcaaa ctccagtgca aactttactt ttcctggtta tcccattcat  2580
299 gtaccagcag gtgtgacact acagactgta tctggtcacc tttataaagt caatgtacca  2640
301 attatggtga cagagacttc tggaagagca ggtattcttc agcatccaat tcagcaagta  2700
303 tttcaacagc ttggccagcc ttcagtaata caaactagtg ttccacaatt gaatccatgg  2760
305 tctcttcaag caactactga aaaatcacag agaattgaaa ccgtgctaca gcaacccgca  2820
307 attctacctt ctgggccagt agataggaaa cacttagaaa atgccaccag tgatatactt  2880
309 gtatctcctg gaaatgagca taaaatcgtg cctgaagctt tgttgtgtca tcaggaaagt  2940
311 tctcactata tcagtcttcc aggtgttgta ttttctccac aggtctctca aacaaattct  3000
313 gatgtggagt cagtgctcag tggttcagct agcatggctc aaaatctgca tgatgagtcc  3060
315 ctctccacaa gccctcatgg ggctctccac cagcacgtga ctgatattca gcttcatatt  3120
317 cttaaaaata ggatgtatgg atgtgattct gtaaagcaac caagaaatat agaggaaccc  3180
319 agcaacatac ctgtatcaga gaaggattct aattctcagg tggatttaag cattcgggtt  3240
321 actgatgatg atattggtga aataattcaa gtagatggaa gcggtgatac atcttccaat  3300
323 gaagaaatag gaagtacaag agatgcagat gagaatgaat ttctagggaa tattgacggg  3360
325 ggagatctga aggtacctga agaagaagct gacagtattt caaatgagga ttcagccaca  3420
327 aacagtagtg ataatgaaga ccctcaagta aacattgtag aagaggaccc tttaaattct  3480
329 ggagatgatg ttagtgaaca ggatgtgcca gacctgtttg acacggataa tgtaattgtc  3540
331 tgtcagtatg ataagattca tcgaagcaag aacaaatgga aattctattt gaaagatggt  3600
333 gttatgtgtt ttggagggag agactatgta tttgcaaaag ccattggtga tgcagagtgg  3660
335 taaaccttgt gagctcagta catctatttt gtgaacatca gttggactat attgcatatt  3720
337 gtgaattcat ttttattttg aatatagtcc agcacagagc tgttcaaatt tttagttcac  3780
339 tgtatggaat ttaataaaat tataattcag atgcagatac aatt                    3824
342 <210> SEQ ID NO: 4
343 <211> LENGTH: 1181
344 <212> TYPE: PRT
345 <213> ORGANISM: Homo sapien
347 <400> SEQUENCE: 4
349 Met Cys Ser Thr Asn Pro Gly Lys Trp Val Thr Phe Asp Asp Asp Pro
350 1               5                   10                  15
353 Ala Val Gln Ser Ser Gln Lys Ser Lys Asn Phe Pro Leu Glu Asn Gln
354             20                  25                  30
357 Gly Val Cys Arg Pro Asn Gly Leu Lys Leu Asn Pro Pro Gly Leu Arg
358         35                  40                  45
361 Glu Phe Pro Ser Gly Ser Ser Ser Thr Ser Ser Thr Pro Leu Ser Ser
```

```
369 Leu Ser Thr Pro Thr Lys Asp Phe Pro Gly Phe Pro Gly Ile Pro Lys
370               85              90              95
373 Ala Gly Thr His Val Leu Tyr Pro Ile Pro Glu Ser Ser Ser Asp Ser
374           100             105             110
377 Pro Leu Ala Ile Ser Gly Gly Glu Ser Ser Leu Leu Pro Thr Arg Pro
378           115             120             125
381 Thr Cys Leu Ser His Ala Leu Leu Pro Ser Asp His Ser Cys Thr His
382       130             135             140
385 Pro Thr Pro Lys Val Gly Leu Pro Asp Glu Val Asn Pro Gln Gln Ala
386 145             150             155             160
389 Glu Ser Leu Gly Phe Gln Ser Asp Asp Leu Pro Gln Phe Gln Tyr Phe
390           165             170             175
393 Arg Glu Asp Cys Ala Phe Ser Ser Pro Phe Arg Lys Asp Glu Gly Ser
394           180             185             190
397 Asp Ser His Phe Thr Leu Asp Pro Pro Gly Ser Lys Lys Met Phe Ser
398           195             200             205
401 Ser Arg Asn Lys Glu Met Pro Ile Asp Gln Lys Ser Leu Asn Lys Cys
402       210             215             220
405 Ser Leu Asn Tyr Ile Cys Glu Lys Leu Glu His Leu Gln Ser Ala Glu
406 225             230             235             240
409 Asn Gln Asp Ser Leu Arg Ser Leu Ser Met His Cys Leu Cys Ala Glu
410           245             250             255
413 Glu Asn Ala Ser Ser Phe Val Pro His Thr Leu Phe Arg Ser Gln Pro
414           260             265             270
417 Lys Ser Gly Trp Ser Phe Met Leu Arg Ile Pro Glu Lys Lys Asn Met
418           275             280             285
421 Met Ser Ser Arg Gln Trp Gly Pro Ile Phe Leu Lys Val Leu Pro Gly
422       290             295             300
425 Gly Ile Leu Gln Met Tyr Tyr Glu Gln Gly Leu Glu Lys Pro Phe Lys
426 305             310             315             320
429 Glu Ile Gln Leu Asp Pro Tyr Cys Arg Leu Ser Glu Pro Lys Val Glu
430           325             330             335
433 Asn Phe Ser Val Ala Gly Lys Ile His Thr Val Lys Ile Glu His Val
434           340             345             350
437 Ser Tyr Thr Glu Lys Arg Lys His Ser Lys Thr Glu Val Val His Glu
438           355             360             365
441 Pro Asp Ile Glu Gln Met Leu Lys Leu Gly Ser Thr Ser Tyr His Asp
442           370             375             380
445 Phe Leu Asp Phe Leu Thr Thr Val Glu Glu Glu Leu Met Lys Leu Pro
446 385             390             395             400
449 Ala Val Ser Lys Pro Lys Lys Asn Tyr Glu Glu Gln Glu Ile Ser Leu
450           405             410             415
453 Glu Ile Val Asp Asn Phe Trp Gly Lys Val Thr Lys Glu Gly Lys Phe
454           420             425             430
457 Val Glu Ser Ala Val Ile Thr Gln Ile Tyr Cys Leu Cys Phe Val Asn
458           435             440             445
```

VERIFICATION SUMMARY   DATE: 06/21/2002
PATENT APPLICATION: US/09/703,809A   TIME: 11:25:19

Input Set : A:\119941-1083 Seq ID Listing.ST25.txt
Output Set: N:\CRF3\06212002\I703809A.raw